10/523221
DT12 Rec'd PCT/PTO 31 JAN 2005

PERTURBATION EFFECT RECOGNITION IN A BIOLOGICAL SYSTEM

Field of the invention

The present invention relates a methods and devices for the machine recognition and rank ordering of perturbation effects on biological systems, and a method of doing business using related to said methods and devices.

Prior art

In general terms, there has been an unmet need for screening systems that can rank order the biological effects (e.g. potency, toxicity, antagonism, inhibition etc.) of new and established chemical entities (henceforth called "drug candidates"). The ability to rank drug candidates with respect to one or more desirable/undesirable features is central to increasing the efficiency of secondary screening, irrespective of assay type. For each drug development programme, many hundreds of decisions have to be taken about halting or permitting continuation of a lead. The biological assay data on which these decisions are based is often disparate, indirect and of variable quality. Many leads are unavoidably lost through inability to differentiate mixed beneficial and adverse characteristics. This high rate of attrition has become a serious problem in drug discovery, undermining the advantages gained by application of ultra-high throughput screening (uHTS) technologies. The desire to lower the 'rate of attrition' has become something of a cliché at international drug discovery meetings, yet there is still no screening technology that can convincingly claim to provide a comprehensive solution to the problem. Additionally there is an interest in measuring the effect of chemical and physical perturbations on biological systems in order to aid the identification of agents causing changes in biological systems. For example if exposure to a perturbation X typically cause a change Y in a biological system, then if an example of the same biological system exhibits a change Y then it could be indicative that the biological system has been exposed to perturbation X.

Object of the invention

The object of the present invention is to overcome some of the problems in the prior art. This is achieved by some or all of the systems, method and methods of doing business having the features of the independent claims 1, 8 and 12.

Summary of the invention

The present invention provides method and systems able to perform some or all of the present functions:

(i) systematize execution of multiple pharmacological experiments on a given physiologically-responding cells/tissue model;

(ii) recognize and score the outcome of the experiments within finite alternative responses (as predicted by an hypothesis set specific to the biological response);

(iii) use the scores (or temporal profile) of responses across all corresponding tests to rank order responses;

(iv) feedback to the user conclusions based on the input rank order criteria (i.e. desired / undesired characteristics, features, data, behaviour).

In a preferred embodiment of the present invention, a substantially automated system and a method are provided which are suitable for monitoring complex biological responses to drugs and/or native physiological ligands, capturing and displaying relative physiological changes induced by specified concentrations of drug/ligand, and comparing the 'captured' response profile (over a defined time-course) to previous iterations of the response (at different concentrations and with different drugs /ligands).

Brief Description of the Figures and Tables

Figure 1 shows schematically a first embodiment of a drug effect ranking system in accordance with the present invention;

Figure 2 shows a flow diagram for an embodiment of an electrophysiological method in accordance with the present invention;

Figure 3a) shows a post-drug signal representing an electrophysiological response of a sample to adrenaline;

Figure 3b) shows a baseline (pre-drug) signal representing the sample of fig 3a) before treating with adrenaline;

Figure 3c) shows a rule representing the drug response result for adrenaline;

Table 1 shows a list of drug response parameters extracted for drugs tested in an experiment in accordance with the present invention; and,

Tables 2a) to 2c) display recognition results for drugs tested in the experiment.

Detailed description of embodiments of the present invention.

The present invention relates to effect ranking systems for use with biological systems, such as cells, tissue, organs and whole organisms. An "effect" in the sense of the present invention is the response over time of the biological system to a perturbation, where a perturbation could be caused by a change in its physical environment (e.g. changes in temperature, humidity, etc.), its chemical environment (e.g. changes in the concentrations of nutrients, waste products, the absence or presence of drugs or hormones, etc), its health, (e.g. ageing, disease, illness, genetic modifications, etc). Such effect ranking systems may comprise some or all of the functions of (a) discovering, preferably automatically, effect rules for each specific perturbation (distinguishing those system measurements giving useful information from those which do not) thereby reducing the number of measurement needed and reducing the effort required to analyse these measurements; (b) describing (e.g. encoding, vectorizing) the biological effects of perturbations in an efficient manner, thereby vastly reducing the size of the datasets stored; (c) representing an 'idealized' perturbation-effect signal curve corresponding to a desired response; (d) comparing multiple actual perturbation-effect signal curves produced with different perturbations to previous results; and (e) rank-ordering perturbation responses with respect to idealized and previous responses according to how closely they evoke the desired response.

An example of such an effect-ranking system could be a drug candidate ranking system. In order to find out if a new chemical would be a suitable candidate to replace a known drug of proven effect, then the drug candidate ranking system would a) discover, preferably automatically, the effect rules for the known drug (i.e. what effect the drug caused on the sample); (b) describe the biological effects of the known drug in an efficient manner; (c) represent a user-defined 'idealized' drug-effect signal curve corresponding to a desired response; (d) test multiple drug candidates, record the drug candidate-effect signal curves, and compare them against the idealized drug-effect signal curve; and (e) rank-order the candidate drug responses with respect to the idealized responses according to how closely they evoke the user-defined idealized response.

Figure 1 shows schematically a first embodiment of a drug effect ranking system in accordance with the present invention. System 1 comprises sample holding and testing means 3 connected by a data buss 5 to a computing device such as a personal computer 7. Computer 7 has operator controlled input means such as a keyboard 9 and pointing device, e.g. mouse

11, and information output means such as a screen 13, printer 15, and data storage means such as a hard disk 17.

Sample holding and testing means 3 comprises a device, such as a clamp 19 able to hold a sample 21, such as a cell, or a plurality of cells, or a piece of tissue, or an organ, or an animal or plant. Sample holding and testing means 3 is optionally provided with an environmental control means 23 optionally comprising a source of fluid 24 and/or temperature controlling means 25 (shown schematically as a fan, but actually comprising any necessary environment controlling means such as heating/cooling elements, humidity controlling elements, etc.) in order to keep the sample alive and at a controlled temperature. Fluid from source or fluid 24 can, for example, be a buffer or nutritional fluid or growth media. Environmental control means 23 may optionally feature other life maintaining services such as means for removing waste products from the cell or tissue. Sample holding and testing means 3 is provided with drug dosing means 27 for applying a drug in a controlled manner to the sample 21. Drug dosing means may, for example, comprise a robot arm 27 able to add to the sample 21 a specific volume of a drug 28m from a platter 29 containing one or more drugs 28a-28n. Sample holding and testing means 3 is provided with sample-response detecting means such as one or more electrodes 31 in contact with the sample 21 and connectable to data buss 5. Data buss is 5 is provided with a signal conditioning means 33 such as preamplifier 35 and amplifier and filter circuits 37 in order to adapt the signals from the electrodes 31 so that they are compatible with the computer 7.

Computer 7 is provided with software programs for controlling the sample holding and testing means 3, the environmental control means 23, drug dosing means 27, and preferably further software for the analysis of the data received from the data buss. Computer 7 is further provided with the usual hardware necessary for executing the software and for storing data. Computer 7 may optionally be provided with means for interoperably accessing data from, and preferably also adding data to, a database of drug response results, said database being stored in any suitable location such as on the computer hard disk or a memory medium such as a disk (e.g. floppy, CD, minidisk, etc.) or memory storage drive, or on a local network or on a server connectable to the Internet.

A flow diagram for a method for using a drug effect ranking system in accordance with the present invention is shown in figure 2. The method comprises the following steps:

Step 110: A sample 21, e.g. a cell, cell network or a piece of tissue, etc., is placed in clamp 19 in sample holding and testing means 3.

Step 120: The sample response detecting means (e.g. electrodes 31) are moved into contact with the sample 21.

Step 130: The natural or evoked biological pre-elicitor signals of the sample 21 (e.g. the signals before the addition of any drug) are detected through sample response detecting means 31, amplified and filtered as necessary by signal conditioning means 33, and sent to the computer 7.

Step 140: These pre-elicitor signals are stored according to a predetermined protocol, e.g. a predetermined number of cycles of sample activity are recorded.

Step 150: The computer 7 commands drug dosing means 27 to dose the sample 21 with a known quantity of a perturbation e.g. drug being tested e.g. drug 28a, and the resulting signals, the perturbed (e.g. drugged) response signals, of the sample are stored according to a predetermined protocol, preferably the same predetermined protocol as used in step 140.

Step 160: The sample 21 may be returned to its natural state by removing the source of perturbation (e.g. washing away the drug) and then the perturbation (e.g. drug) testing procedure repeated on the sample 21 with the same perturbation (e.g. drug), possibly at a different concentration, or with a different perturbation (e.g. a different drug) being tested.

Step 170: Alternatively the sample 21 may be discarded.

Step 180: The stored signals from each test are processed in software.

Step 190: If the purpose of the test is to find the sample response to a known perturbation (e.g. a drug) then the differences between the pre-perturbation (e.g. pre-drug) signals and the perturbed (e.g. drugged) response signals are calculated by the software and analysed, as explained in more detail below, in order to produce a perturbation response result which is characteristic of the perturbation causing the result.

Step 200: Each perturbation response result is then stored in computer memory, preferably in a database where each known drug and its drug response result are stored with each other.

Step 210: If the purpose of the test is find out if the drug being tested produces a electrophysiological response which is similar to a known drug then, after the drug response result for the drug being tested has been produced, it is compared against the electrophysiological drug response results stored in the database.

Step 220: Response matching software, e.g. software able to identify and match points on the pre-drug and post drug signals which relate to the same type of biological event, orders the drug response results from step 210 and preferably produces a table in which the drug response result for the known drug which most closely matches the drug response result for the drug being tested or closely matches to parameters extracted from a desired theoretical or real result is ranked highest, the drug response result for the known drug which next most closely matches the drug response result for the drug being tested is ranked second highest, and so on.

Preferably the drug effect ranking system is provided with auto-test and/or calibration means. This is because of the complex nature of biological signals makes it difficult for a user to know if a signal is a product or biological activity or a fault in the device. (i.e. the user may wish confirmation that the instrument is working to specification). Therefore, the system is preferably provided with an auto-test and/or calibration routine that can be accessed by an operator e.g. via a dialog box on the computer screen 13, or automatically, for example, as part of a warm-up routine for the system or if the system detects a fault e.g. via a call from a system auto-test. This could comprise a cell activity simulator that would embody a 'dummy' sample that is able to inject into the system electrodes 31 electrical signals of similar magnitude, character and (if relevant) periodicity/amplitude of 'authentic' biological samples (i.e. as if they were actually present on the device type). This may be achieved by playback of a sampled waveform or electrical characteristic, preferably looped end-to-beginning so that the data output is continuous. This type of recording may be made from a genuine sensor under ideal conditions, or synthesised from component waveforms, or a combination of these. The 'dummy' sample may be connected to all the recording channels on the data buss to facilitate signal checking. The cell activity simulator preferably has sufficient features to (i) assist instrument calibration, (ii) aid quality control, (iii) reduce spurious fault reporting (e.g. where a customers' cells are not active but the customer thinks/reports the instrument is not working to specification), (iv) provide internal diagnostic/ amplification channel fault-testing

routines (as above), (v) training for instrument data capture and processing and (vi) for instrument capacity demonstrations during 'white coat selling' or at trade shows, where cells (especially a wide variety of cellular phenotypes) are not readily available. The signals generated from cell activity simulators preferably are capable of feature extraction and processing like 'live' signals. The availability of a 'simulator' routine means that instrument trainees may gain experience of data processing without having to wait for the availability of responding cells.

As recording the effect of drugs on samples such as living cells, tissues, organs and the like requires drugs to be dispensed under near-physiological conditions in growth media, as opposed to a simple physiological buffer or saline, and to provide accurate information the system also requires the sample to be maintained within specified physical limits on the sample holding and testing means 3 prior to, during and between electrical activity readings. Media or media plus drug solutions need to be pre-mixed, pre-warmed and added to cells that might exhibit lability to osmotic, pH shift or mechanical shock. Therefore environmental control means 23 may control the heating of fluid dispensing system 24, and may also be provided with sample-contacting sensors (not shown) for collecting data relevant to cell health (e.g. but not exclusively pH / osmolarity / temperature). The maintenance of physiological responsiveness when adding media containing drugs to the cell and sensor requires a special sequence of operations. These steps must deliver minimum non-specific 'shock' to the cells (i.e. the only desired change to the cells would be the presence of the drug) – this reduction of 'shock' can be achieved by including universal 'pre-incubation' of the cells with medium that contains only the solvent vehicle for the drug under test. This pre-incubation requirement satisfies the experimenter that any physiological effects observed were entirely due to the compound and not the solvent/vehicle. In order to avoid mechanical artefacts, the addition of media containing drugs is preferably performed relatively slowly at a point of delivery away from the sensing electrodes 31. The needle dispenser operation in the figures below, sequence of operation and relative placement within the well of the dispenser/aspirator head can to be optimised for the fulfilment of these requirements. Additionally, an important requirement in physiological drug profiling is the maintenance of cell health / viability / responsivity especially during long time-course experiments. Thus environmental control means is preferably programmable to act in 'life-support' mode, where

drugs are not added, but cell growth media is removed / added at regular intervals. This 'cell maintenance' mode of operation may perform some or all of the independent functions of:

1.Removing waste materials from actively-respiring cells/tissue;

2.Supplying replacement 'fresh' oxygenated, pH-controlled media;

3. Providing conditions that are used as data sources for multiple 'baseline' activity measurements – i.e. recording activity of cells that have <u>not</u> received drug;

4. Providing conditions that are used as data sources for mechanical perturbance control activity (i.e. a system check that the per se physical force accompanying injection of drugs in media onto a cell preparation has not perturbed the electrophysiological signal); and,

5. Providing baseline data for 'vehicle alone' controls (i.e. for reporting any activity changes that might arise due to non-specific physiological effects of the solvent which the compounds are dissolved in).

As conditioning the biological signals involves large amplifications of the biological signals it is important to suppress the generation or transmission of electrical noise and any other form of interference (e.g. mechanical vibration, radiofrequency) in order to maximise signal-noise ratio. This is preferably achieved by providing isolation mechanisms from mains-borne interference, and preferably using an instrument-dedicated earth (e.g. an efficient earth dissipation route / grounding spike). A fine-mesh, earthed Faraday cage is preferably used around the sample. Preferably the electrical power supply paths to different system components are optimally distributed to achieve low-noise.

Different cellular substrates have different requirements for environmental control (actively respiring, spontaneously active neurons or cardiomyocytes will be much more fastidious in their environmental, nutritional and waste-removal requirements than, say Xenopus oocytes). Therefore there is occasionally a need to have a local 'microenvironment sustaining mechanism' to control humidity in order to reduce evaporative loss of moisture and there may be requirements for keeping the degree of oxygenation and pH control within acceptable (physiological) limits. These requirements may be met by providing environment control by using a thermostatted chamber. To maintain visibility of the experimental samples the chamber is preferably made of a transparent material (e.g. Perspex™ or the like). The facility to connect optional thermostatic heater units and provide a supply of CO_2 may be optionally provided.

In the systems and method of the present invention reliable methods of variable selection and prioritization are used to provide reproducible rules for identification of drug biological effects from raw biological data.Matrix calculations can then be performed to identify drug-specific functional parameter changes and was capable of 'neighbouring' (similarity analysis) and rank-ordering. A general understanding of relevance when developing screening platforms to measure any biological response is that one can easily populate databases with parameters that are essentially bogus - i.e. they might vary with drug application but ultimately are of no predictive value in understanding the biological action of a drug. Previous attempts to define the 'genuine' parameters, e.g. in electrophysiology, have been rather *ad hoc* – however the present invention provides methods that can discover which parameters are most important to study and encapsulate these in a predictive machine-readable matrix.

A comparator program can perform 'supervised learning', i.e. it can learn relationships from examples the operator provides using genetic programming or evaluation computing techniques or the like. The operator can then use the learned relationships to interpret new data of similar characteristics. Alternatively, comparator program can be used to search for some form of relationship, such as 'which are the important variables' – For example, *which* cDNA sequences predict susceptibility to a tissue-selective disease? It has strengths in discovering why some samples are observed to be different from others. Because a comparator program handles regression problems in an unusual way – by ranking - it can predict the (continuous) value of a Y variable in both non-linear and linear relationships. A frequent use of a comparator program in drug discovery is in mining data – e.g. what structure is 'good' for druggability - by exploring the characteristics of data, it can find the important variables.

Experimental results

An example of an experiment using the above system and method was as follows:

18 mouse cardiomyocyte preparations were tested. Spontaneously- active cardiomyocytes were made from approximately 80 pregnant Balb-c mice with average yield of 0.5-1 x 10^6 cells per preparation (average ~1 x 10^4 cells per E13 embryo). Approx. 5-10 multielectrode arrays (average 6/prep) were seeded with 1 x 10^5 cells per preparation. Some variation in the quality of cells grown from these native sources was encountered. Approximately 40% of seeded multielectrode arrays were considered to give 'useable' spontaneous electrical activity before compound addition. The number of electrodes/channels on multielectrode arrays

giving adequate results varied from 10-60%. The newer the multielectrode arrays the better the pre-drug 'baseline' activities. Recycling of multielectrode arrays was generally problematic. The study revealed that much variability in recordings was due to the proximity of electrically active cells to electrodes (i.e. physical distance affects the shape of the signal). Pre-drug and post-drug signals from 30 different compounds were collected with 'repeat' data on 5 of these compounds.

Continuous raw data records (consisting of many cardiac cell electrical cycle responses pre-drug and post-drug) produced on the multielectrode arrays system were ported to Microsoft Excel ™ by using a data extraction tool i.e. these relatively large data files were 'cropped' manually to encompass one or two full pre-drug and post-drug electrical response cycle (see figures 3a) and 3b) for a post drug and baseline response cycles for adrenaline) and these smaller files (albeit still several Mb) were imported into MATLAB ™ (from The Mathworks, Inc. Natick, Mass. USA)) for feature extraction i.e. 17 potentially informative parameters (i.e. those that one would normally be able to predict drug action using knowledge of cardiac cell electrophysiology and multielectrode array principles) were extracted by a program specially written to extract the height, depths and time parameters forming the 17-parameter set for 22 drug effects (see table 1). The selection of which parameters to select from the raw data was based on the ability to identify the same biological event occurring before and after drug addition. Parameters for all of the sample sets (analogised as 'spectra') were calculated, with parameters after drug addition being subtracted from the same parameters before drug addition. This operation gave a 'difference' or 'delta' reading that was mapped to actual biological events. There was no prospect of seeing which of the possibilities was the 'best' to choose – so this task was reserved for a comparator program. A comparator program is designed to derive simple relationships using a small number of variables in large multivariate data sets – i.e. it only 'picks' those variables that are seen as significant. Typical comparator programs (e.g. gemax-bio, Aber Genomic, Aberystwyth, UK) favour the generation of simple rules first and then add complexity, selecting by using evolutionary criteria). This feature is of some importance when analysing high-dimensional electrophysiology data sets. Overall, a comparator program learns from examples that the operator provides, giving rule-based answers, using only those variables in the data set that are important in relation to the specific problem. Unlike techniques such as neural networks (where the path to a solution is usually not visible to the operator) the supervised learning method in comparator program states any

derived solution (rule) in plain text. Once a comparator program had discovered any rules for a set of data, they are generally human-interpretable and exportable in ways that do not require the comparator program itself (see figure 3c for the data rule for adrenaline).

The intermediate 'biological event mapping' step was performed in order to avoid misleading phase errors that are introduced when action potential data were subtracted automatically from a single matched point (e.g. the peak of the Na^+-dependent depolarization). This step was necessary, as it was clear that as well as drug actions introducing Y-axis (signal amplitude) variance in the data, they also introduced X-axis variance (i.e. lengthened and shortened time values of excitable phases). The net effect of this X- and Y-axis covariance was that simple arithmetic subtractions of pre-post-drug would not work consistently. In the present invention. In other words, the underlying biological processes causing a event (and its subsequent signal) were identified ("mapped") and the biological events were used to match the signals in the recorded records. Event mapping thus ensured that both X-axis and Y-axis parameter value changes induced by drugs were recorded.

A set of values that represented changes (or 'delta') effects of drugs on the biological system (i.e. pre-drug versus post-drug) were inputted into the comparator program. This was determined as an important observation, as the system and method is intended to rank order the 'drug effect' and therefore there was more value in entering the 'delta' signal. This avoided the problem that if raw data alone was used to profile drug responses, then complex non-linear time-base shifts in waveform shapes would mean comparisons were not being made 'like with like'.

The 'drug effect profiling' principle can not be couched in terms of a simple 'is this result in one class or another' which comparator program is well-able to solve. Drug effect requirements are seen as more complex: there is a set of native or heterologous screening targets expressed in a cellular system; and also a set of assays that are technically appropriate for reporting the biological responses within each cell system. Therefore it has to be re-considered how the comparator program's rule-based system alone could be applied. Practically, many (potentially hundreds of thousands) biological response-class 'rules' can be run encapsulated in a single script. It is possible to apply <u>many</u> 'rules' describing previous response classes to <u>one</u> compound-induced response at a time. For example, a system having

say 9,876 response models (profiles) could assign an new drug response to a certain 'class 235' classification by virtue of it having a high 'response model 235' score (...'is it response class 235 or not, decided by a ratio or probability function...). Numerical ranking of score to other response models (i.e. 0001-0234 and 0236-9876) would enable rank ordering effects. Each of the thousands of rules can be applied automatically, although in the present experiment, only 22 such 'drug effect' rules were available.

Selected primary data produced the drug-effect models (and thereby rules, reproduced in the Results section). For any new drug test response result (or a theoretical ideal test result) it is interesting to discover which drug response it was 'most like'. For this reason, the comparator program may be run against every single drug response type and the correspondence with the drug response types ranked.

The application of the comparator program to generate drug-response-specific rules was successful. An analytical-based technique was then applied in order to recognise drug-specific responses and physically rank-order them. This approach consisted of a numeric application to the data incorporating some elements of PCA (principal component analysis). In brief, the 17 parameters extracted from the raw data were tabulated in a data array with rows representing n compounds. For convenience of description, these parameter sets can be considered (analogised) as a 'spectral' data array of components as follows:

Compound	Par 1	Par 2	Par 3	Par 4	Par 5	Par 6	Par 7	Par 8	Par 9	Par 10	Par 11	Par 12	Par 13	Par 14	Par 15	Par 16	Par 17
A																	
B																	
C																	
n																	

The assertion then used was that the 'spectrum' of an unknown (or designed or optimised or desired) biological response to compound X (S_X) is equal to the sum amount of all 'component-of-total' response parameters (C), i.e.

$S_X = X_1C_1 + X_2C_2 + X_3C_3 + X_4C_4 + X_5C_5 + X_6C_6 \ldots$

In a matrix representation, all parameters are rows and values are columns. Different biological responses can be represented by 'a linear mix' of simpler responses that 'add-up' to a more complex response in a way that is dependent on the concentrations of the components (this reasoning was borrowed by analogy from behaviour of additive spectra). Thus it was assumed that typical biological response results are equivalent to a mixture of simpler responses. Additional information can be applied about presence or absence (and concentration) of drugs in a matrix calculation. In essence, if one knows what all of the response mixture compositions and all of the drug concentrations were, one can derive what the sub-components should be like. A potential problem with this approach occurs when there is only one example available for each drug response. In effect, all this is capable of producing is an identity matrix, which is not useful. However, the multielectrode arrays data set also comprised some repeat data, so this was used to construct a new set of matrices named the results matrix, fingerprint matrix and concentration matrix, as shown below.

Observed Feature Matrix (O) (Results Matrix)

	Feature 1	*Feature 2*	*Feature 3*	*Feature p*
Drug 1	F_{11}	F_{12}	F_{13}	F_{1p}
Drug 2	F_{21}	F_{22}	F_{23}	F_{2p}
Drug 3	F_{31}	F_{32}	F_{33}	F_{3p}
Drug 4	F_{41}	F_{42}	F_{43}	F_{4p}
...				
Measurement m	F_{m1}	F_{m2}	F_{m3}	F_{mp}

Drug Concentration Matrix (C)

	Drug 1	*Drug 2*	*Drug 3*	*Drug n*
Measurement 1	1	0	0	0
Measurement 2	0	1	0	0
Measurement 3	0	0	1	0

Measurement 4	0	0	0	0
...				
Measurement m-3	1	0	0	0
Measurement m-2	1	0	0	0
Measurement m-1	0	0	1	0
Measurement m	0	0	1	0

A matrix K can be defined which contains the idealised drug response parameters for the drug response fingerprint matrix:

	Feature 1	Feature 2	Feature 3	Feature p
Drug 1	I_{11}	I_{12}	I_{13}	I_{1p}
Drug 2	I_{21}	I_{22}	I_{23}	I_{2p}
Drug 3	I_{31}	I_{32}	I_{33}	I_{3p}
Drug 4	I_{41}	I_{42}	I_{43}	I_{4p}
...				
Drug n	I_{n1}	I_{n2}	I_{n3}	I_{np}

Using Matrix formulation comes the relationship $O = CK'$
Using least squares analysis in MATLAB™ K can be calculated knowing O and C
Once K is established the features of the response to an unknown drug R_u can be used to calculate the similarity of drug response to the known drug responses S using the relationship $R_u = SK'$

Matrix multiplication was performed in MATLAB™. A 'spectrum' of parameters can consist of 'real' results (i.e. experimentally-derived and extracted as described earlier with respect to the comparator program) as well as 'idealised' response parameters that can be input into the drug profiling system as the 'target' for a secondary drug screen result. Calculation using matrices 'concentration' and 'results' can populate the response fingerprint matrix. Any row can be made to represent the ideal response values of the biological system for any particular drug. Application of such matrix calculations to data produced on the multielectrode arrays rig showed that results for two different applications of the same drug did *not* give exactly the

same results. The implication of this finding is that when the parameters are calculated, MATLAB™ does not pick out exactly the same values as previously, but rather an averaged/least squares best fit is found. Thus the 'fingerprint' matrix is a representation of how the drug is expected to work. The final and most crucial step is where one takes an unknown spectrum of parameters (or an operator-defined spectrum inputted by an operator via an input interface) and use the existing matrices to calculate how much of each separate drug response spectra (row 1, row 2, row 3, row 4....) is required to produce the unknown or operator-defined spectrum. As experimental results become available, further matrix calculations can evaluate how 'far' they are from idealised values and in effect how the result could be reconstituted from a 'mix' of the various matrix values. When input procedures were performed for 'real' fingerprint data produced on the multielectrode arrays rig, 12 fingerprints from a total of 21 gave results that indicated recognition of previously-logged drug effect rows (i.e. 1 values in one column with 0 values in all other positions) – see Tables 2(a) to 2(c). This was tantamount to 100% confidence that the biological response was induced by the specific drugs. Other results in the batch gave 80% confidence of recognition (i.e. 0.8 of the response was made up of the 'correct' drug fingerprint components). It is very clear that achievable improvements in data quality (most notably by correction of the electrode 'proximity effect' and the application of improved on-cell electrode designs) would improve prediction. Thus, the method in accordance with the present invention is able to recognise if the output representing the differences between signals from a sample before it is perturbed and signals from said sample after it has been perturbed are the same as, or similar to, an output representing previously tested perturbations.

In essence, the comparator program produced different rules for different drug responses, providing simple rules first and adding complexity if required. Matrix calculations within MATLAB™ used the same 'rule' to identify the response pattern inherent of all drug responses, and all calculations were of equal complexity. However, a major problem in matrix calculation methods when applied to parameters derived from physiological system responses to drugs is selecting which parameters to study. In practice, when matrix fitting was attempted on all 17 parameters the analyses became chaotic, an effect that was interpreted as an attempt to matrix fit noise. Thus, in attempting to fit the data using all 17 parameters, MATLAB™ was taking three parameters of one drug response, two from another, one from a third, minus two from a fourth... etc. This became an 'Achilles heel' of the method as it produced noise. In

other words, large calculations were required in order to solve a small part of the fitting problem. However, the software accepted this behaviour as normal, simply because the matrix calculations had no measure of noise. Crucially, the matrix approach was being followed as a supplementary approach to the comparator program approach that generated rules identifying the most informative parameters for the drug-specific effect(s). In the present invention the problem is solved by informative parameter selection, that is selecting only parameters which provided information about the biological system. If the rules (informative parameters) had been applied as a start point of the matrix calculations, there would have been no 'automatic' attempt to 'fit noise' the software, yielding a generalised efficient recognition method for drug-induced biological responses (e.g. a peak in the signal caused by a sodium channel opening). If one compares all 17 'extractable' parameters from the multielectrode arrays raw data spreadsheets to ones that were actually selected in the comparator program-derived rules, there are only about 8 that were 'informative'. Use of these informative parameters (and ignoring the rest that are measurable but not selected in a genetic program) saves an enormous amount of calculation in getting to the same quality of result.

The comparator program produced different rules for every single drug response – for efficient rank ordering this is actually a problem as it was difficult to make the program look for all the rules at once. The MATLAB™ matrix calculations, on the other hand, were capable of looking at all the results simultaneously (albeit when presented as a self-consistent set i.e. a set wherein each event is compared against the comparable event in the next record). In order to optimise selection of informative parameters, it can be advantageous to have an operator who is familiar with the biology underlying the system monitor the parameter selection/evaluation process. The operator can help select which parameters are informative (i.e. indicative of a drug effect), and reproducible.

While the present invention had been described by examples using continuous electrophysiological data from multielectrode arrays, it is not limited to the use of electrophysiological data. Any appropriate response detecting method and data collected form the response detecting method may conceivably be used in a screening system, for example, for cells, the rate of growth or changes in the expression of proteins may be appropriate responses while in tissues it may be appropriate to record the rate of oxygen consumption or the speed of muscle contractions. It may also be applied to any biological system exhibiting a

change in signals over time, for example, charting the progression of an affliction, disease or illness (or their cures) by comparing records of signals (e.g. analytical profiles, molecular probe localisations, computed parameters, magnetic resonance imaging images, body fluid assays, disease markers, etc.) taken from a patient at time intervals of hours, days, weeks or longer. An example of a drug screening system using the present invention could comprise fluorescence labelled cells. These cells could be immobilised in a cell array (for example, one cell deep on a substrate - analogous to a retinal array) or otherwise positioned so that the position of each cell is known (or each cell is identifiable) such that an imaging device such as a scanner can repeatably return to scan a known cell. A method for testing a drug or drug candidate (or other perturbation) on a labelled cell or group of labelled cells could comprise the following steps:

the position of a first cell(or group of cells) is determined,

the imaging device is positioned to record an image, or images, of the cell (or group of cells),

the cell (or group of cells) is challenged and its response to the challenge imaged and recorded,

the cell (or group of cells) is allowed to return to its unchallenged state (and optionally the imaging device is move to image a second cell or group of cells),

the imaging device is positioned to image the first cell (or group of cells),

the first cell (or group of cells) is perturbed by the addition of a drug or drug candidate,

the first cell (or group of cells) is challenged with the same challenge as used before and its response to the challenge imaged and recorded,

the stored signals are processed as described above.

The present invention may be applied to a method of doing business comprising the step of: providing a database of the biological effects of known drugs which customers may access in order to compare the biological effect of a drug candidate against the biological effect of known drugs.

A further embodiment of a method of doing business in accordance with the present invention comprises the additional step of:

providing software capable of encoding the biological effect of drugs, for example as a vector or a rule or an equation or the like describing a drug response curve or differences between a pre-drug response and a post-drug response. The essential difference between a visualisation

tool and a system in accordance with the present invention is that in the present invention a dataset containing previously recorded perturbation responses can be remotely searched (e.g. via the internet or an intranet) and the system recognises if the perturbation being tested gives a result which is the same as a previously recorded perturbation result. If the result is not identical to a previously recorded perturbation result then the system in accordance with the present invention can rank-order the previously recorded perturbation results according to how closely they match the result for the perturbation being tested.

A further embodiment of a method of doing business in accordance with the present invention comprises the additional step of:
ranking the biological effect of a drug candidate against the biological effect of known drugs.

The above mentioned embodiments are intended to illustrate the present invention and are not intended to limit the scope of protection claimed by the following claims.